Mixture Distributions for Modeling Lead Time Demand in Coordinated Supply Chains*

Barry Cobb

Virginia Military Institute

Alan Johnson

Air Force Institute of Technology

AFCEA Acquisition Research Symposium — May 2014

*Support from grant N00244-13-1-0014 to VMI Research Laboratories, Inc. from the Naval Supply Fleet Logistics Center (NAVSUP) through the Acquisition Research Program at the Naval Postgraduate School is gratefully acknowledged.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **MAY 2014**	2. REPORT TYPE	3. DATES COVERED **00-00-2014 to 00-00-2014**

4. TITLE AND SUBTITLE **Mixture Distributions for Modeling Lead Time Demand in Coordinated Supply Chains**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Virginia Military Institute,Lexington,VA,24450**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
AFCEA 11th Annual Acquisition Research Symposium, 14-15 May 2014, Monterey, CA.

14. ABSTRACT

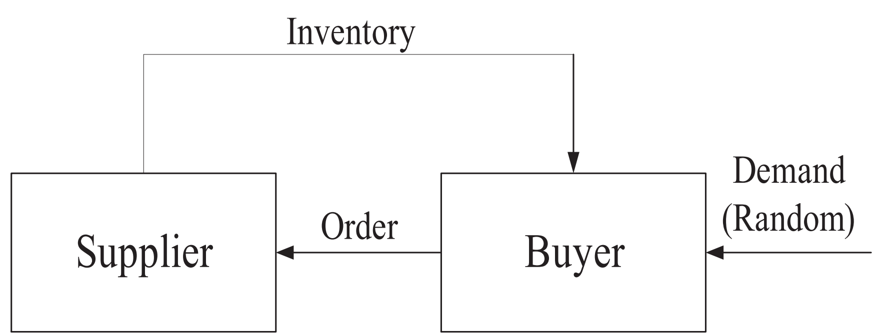
15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**33**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Two-Level Supply Chain

• The buyer randomly receives demand from its customers and places orders of size Q from the supplier when its inventory level drops below reorder point R.



• The supplier receives these orders and ships inventory to the buyer. The supplier orders a quantity from its supplier in integer multiples N of Q.

Two-Level Supply Chain

- Buyer cost function:

$$TC_b(Q,R,V) = (K_b - V)\cdot\frac{Y}{Q} + h_b\cdot(0.5Q + R - E(X)) + \pi\cdot S_R(R)\cdot\frac{Y}{Q}$$

$\mapsto$ Terms are annual ordering, holding, and stockout costs; V is a per order rebate coordination incentive (Cobb and Johnson 2013).

Expected shortage per cycle: $S_R(R) = \int_R^\infty (x-R)\cdot f_X(x)\,dx$

$\mapsto$ f_X is the probability density function (PDF) for lead time demand.

Two-Level Supply Chain

- Supplier cost function:

$$TC_s(Q,N,V) = \left(\frac{K_s}{N} + V\right) \cdot \frac{Y}{Q} + h_s(N-1)0.5Q$$

$\mapsto$ N = integer multiple of buyer's order size Q

$\mapsto$ $\left(\frac{K_s}{N} + V\right) \cdot \frac{Y}{Q}$ = annual ordering costs

$\mapsto$ $h_s(N-1)0.5Q$ = annual holding costs

Lead Time Demand

• LTD follows a compound probability distribution. Suppose L is a random variable for LT and D represents random DPUT. LTD is a random variable X determined as

$$X = D_1 + D_2 + D_3 + \cdots + D_i + \cdots + D_L \ .$$

Therefore, X is a sum of random, independent and identically distributed (i.i.d.) instances of demand. The mean (μ_X) and variance (σ_X^2) of X can be calculated as

$$E(X) = E(L) \cdot E(D) \text{ and } Var(X) = E(L) \cdot Var(D) + [E(D)]^2 \cdot Var(L) \ .$$

Example (Eppen & Martin (1988))

- Daily demand is normally distributed: $D_i \sim N(40, 30)$.

- Lead time (in days) is discrete: $\Omega_L = \{7, 12, 14, 15, 16, 25\}$ each with probability 1/6 ($E(L) = 14.8$; $Var(L) = 29.1$).

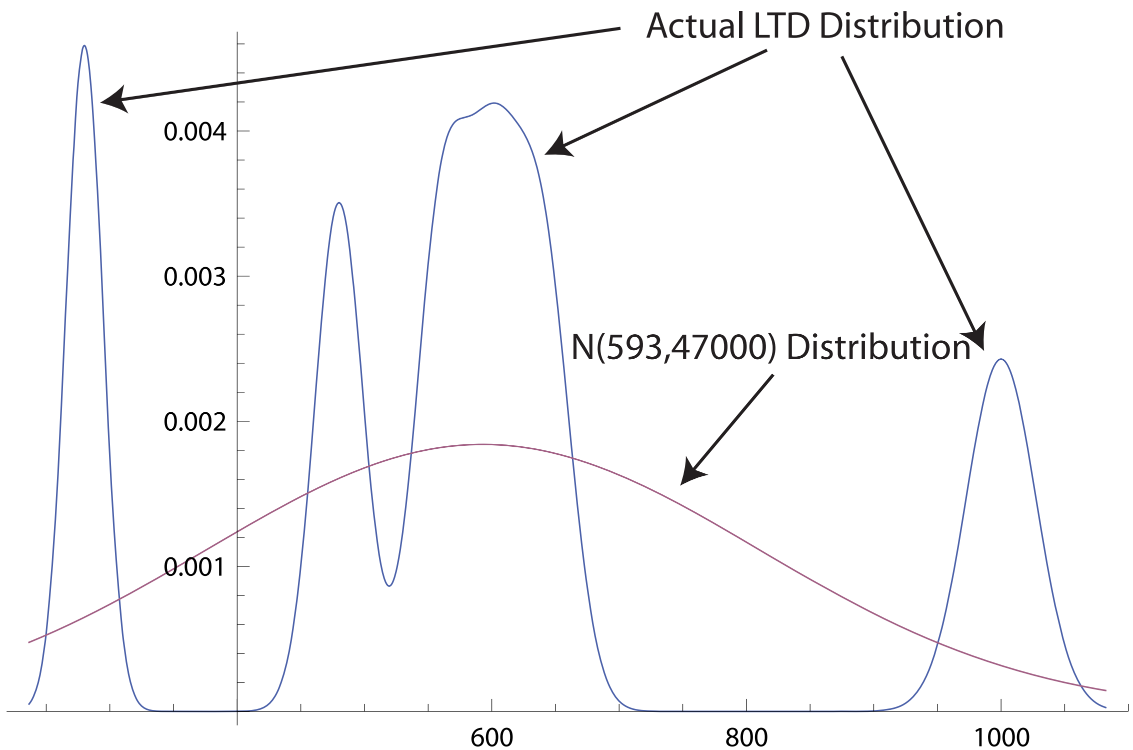
- In this case, $E(X) = \mu_X = 14.8 \cdot 40 \approx 593$ and

$$Var(X) = \sigma_X^2 = E(L) \cdot Var(D) + [E(D)]^2 \cdot Var(L) \approx 47000$$

- All previous methods for setting (Q^*, R^*) and N^* assume LTD is normal; Eppen and Martin (1988) demonstrate calculation of a service level in a single-firm context.

Example (Eppen & Martin (1988))

- The actual LTD distribution, f_X, is a mixture of normal distributions.



Two-Level Supply Chain

- Research issues to be addressed:

1. Analytical solutions for Q^*, R^*, and N^* in the two-level supply chain problem assume (and require) normality.

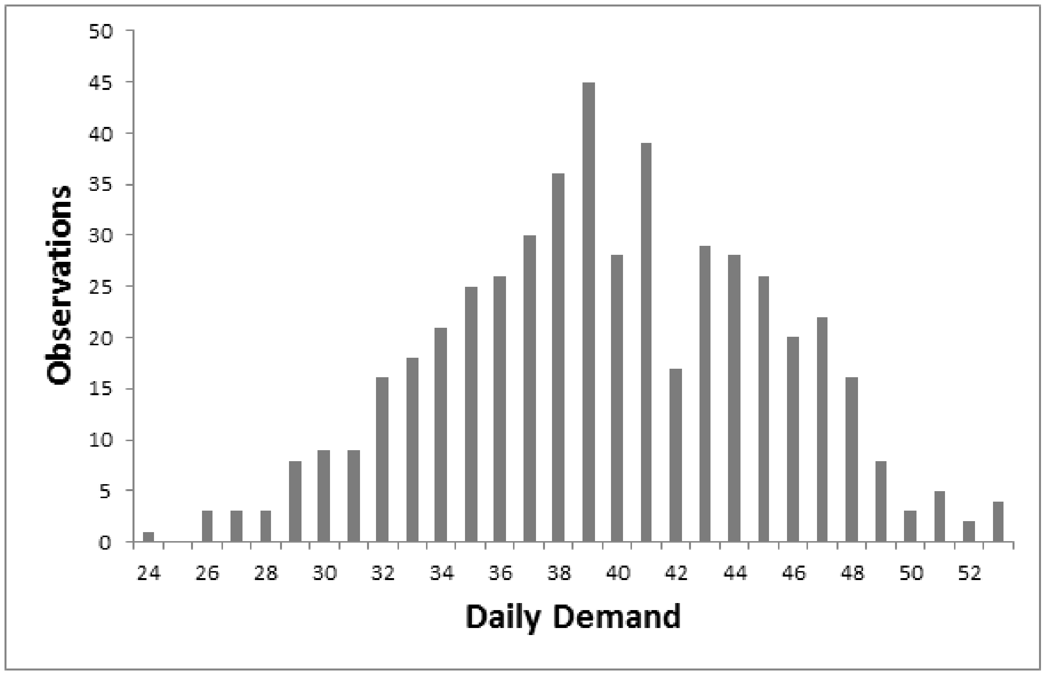
2. Methods for modeling LTD distributions often make unrealistic distributional assumptions.

3. In practice, the actual LT and DPUT distributions are likely unknown – the solution here uses empirical data.

$\mapsto$ For the example, suppose a modest amount of historical data is available on daily demand and lead time on previous orders.

Example (Eppen & Martin (1988))

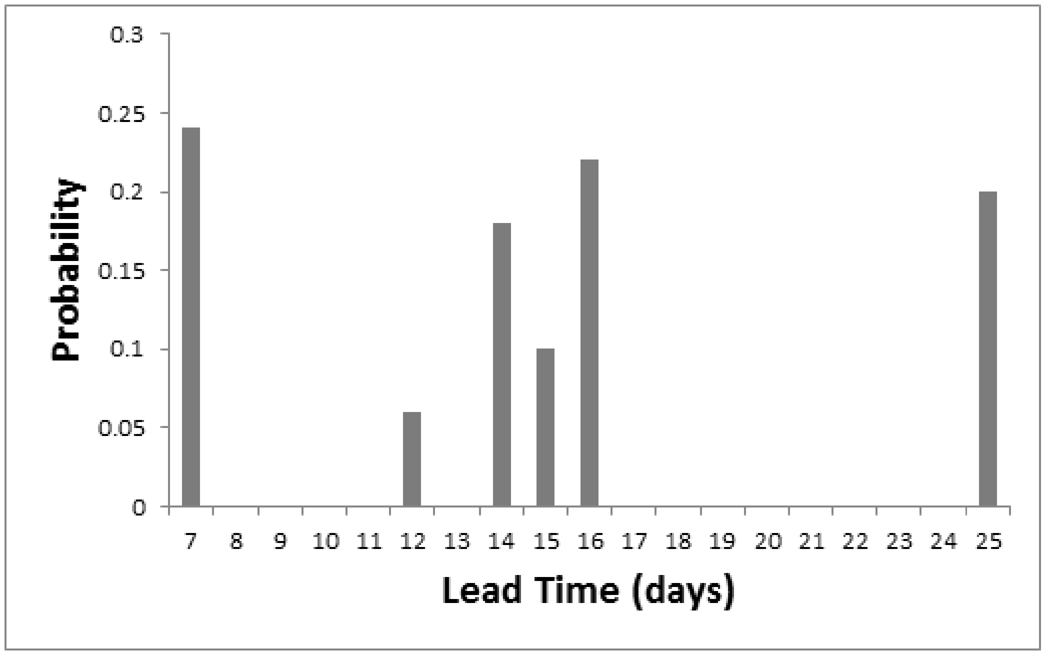
$\mapsto$ 500 observations of daily demand with $\overline{d} = 39.66$ and $s_d^2 = 30.64$.



- This is a random sample from the $N(40, 30)$ distribution.

Example (Eppen & Martin (1988))

$\mapsto$ 50 observations of lead time with $\bar{\ell} = 10.8$ and $s_\ell^2 = 12.52$.



- This is a random sample from the discrete LT distribution.

Constructing the LTD Distribution

$\mapsto$ 500 observations of daily demand with $\overline{d} = 39.66$ and $s_d^2 = 30.64$

40	42	40	47	41	...	40	33
37	45	35	27	42	...	38	32
47	38	47	40	47	...	39	49
47	40	41	36	34	...	38	39
45	49	37	43	38	...	32	37
34	48	45	31	47	...	35	32
31	43	32	47	41	...	36	33
46	46	40	41	38	...	38	47
38	46	33	45	38	...	32	45
33	35	39	41	37	...	39	39
34	50	39	42	39	...	47	
33	39	44	37	32	...	45	
42	41	41	43	39	...	47	
37	46	41	32	34	...	38	

- These are assumed to be i.i.d. observations.

Constructing the LTD Distribution

↦ The most likely empirical LT value is 7 days

	40	42	40	47	41	...	40	33
	37	45	35	27	42	...	38	32
	47	38	47	40	47	...	39	49
	47	40	41	36	34	...	38	39
	45	49	37	43	38	...	32	37
	34	48	45	31	47	...	35	32
	31	43	32	47	41	...	36	33
7-day total	281	305	277	271	290	...	258	255
	46	46	40	41	38	...	38	47
	38	46	33	45	38	...	32	45
	33	35	39	41	37	...	39	39
	34	50	39	42	39	...	47	
	33	39	44	37	32	...	45	
	42	41	41	43	39	...	47	
	37	46	41	32	34	...	38	
7-day total	263	303	277	281	257	...	286	

- Sum daily demand over each 7-day period in the dataset

Constructing the LTD Distribution

$\mapsto$ The lead time demand dataset given LT of 7 days with $N_7 = 71$

281	275	299	260	278	255	287	279	252	302	286
305	254	306	276	289	263	279	278	294	283	
277	263	283	264	283	303	279	280	255	271	
271	272	261	288	298	277	264	281	273	302	
290	280	287	271	261	281	276	276	259	278	
282	297	262	279	276	257	262	266	281	287	
275	295	284	299	258	285	279	236	259	294	

• Fit a mixture of polynomials (MOP) distribution (Shenoy 2012) to this data. This distribution will be the approximate LTD distribution conditional on $L = 7$, or $\widehat{f}_{X|L=7}$.

• Similar distributions, $\widehat{f}_{X|L=\ell}$, will be constructed for $\ell = 12, 14, 15, 16, 25$.

Mixture of Polynomials (MOP)

$\mapsto$ The approximate LTD distribution conditional on $L = 7$, or $\widehat{f}_{X|L=7}$

$$\widehat{f}_{X|L=7}(x) =$$

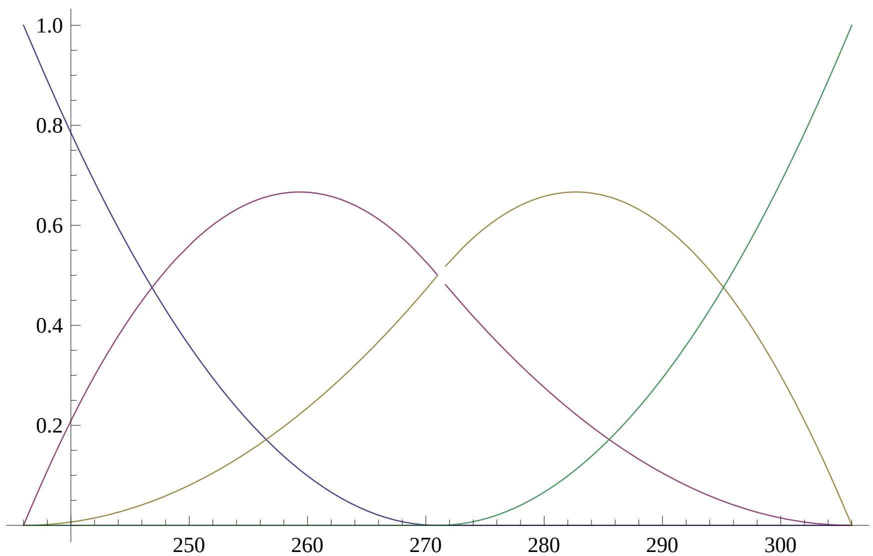
$$\begin{cases} -3.3726 + 0.0241x - 0.000043x^2 & 236 \leq x < 271 \\ 0.7102 - 0.0061x + 0.000013x^2 & 271 \leq x \leq 306 \end{cases}.$$

• The MOP has $n = 2$ pieces and is degree $d = 2$ (or is third order).

$\mapsto$ The MOP was constructed using a linear combination of B-spline functions (Lopez-Cruz et al. 2012).

Mixture of Polynomials (MOP)

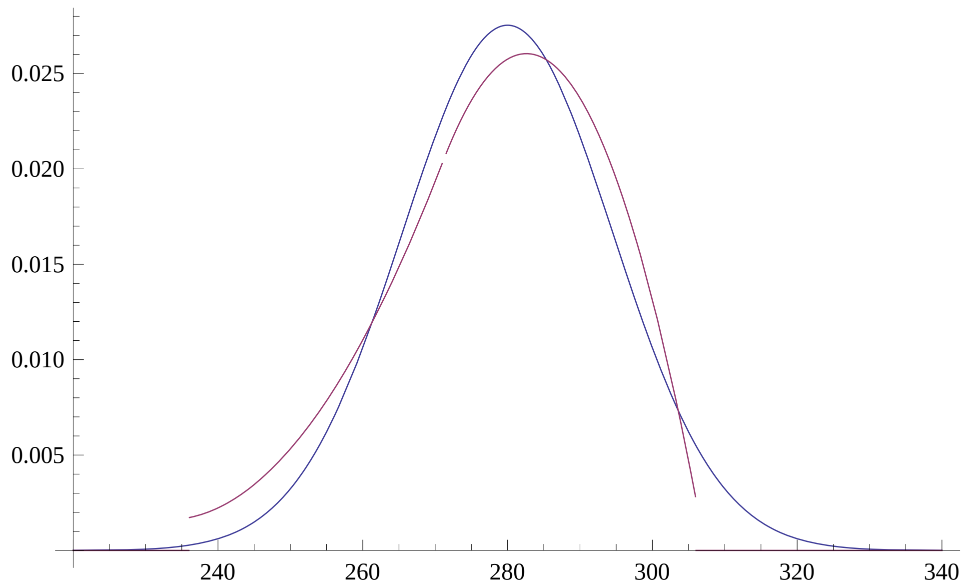
$\mapsto$ $\widehat{f}_{X|L=7}$ is a mixture of four B-spline functions



• The B-splines are defined recursively based on the split points in the domain. Mixing coefficients are determined via maximum likelihood (Zong 2006).

Mixture of Polynomials (MOP)

$\mapsto$ The approximate LTD distribution conditional on $L = 7$, or $\widehat{f}_{X|L=7}$ overlaid on the $N(7 \cdot 40, 7 \cdot 30)$ distribution.



• Recall: the MOP is not fit to the normal PDF, but rather a small sample of data generated from the normal PDF.

Mixture of Polynomials (MOP)

• Selection of d and n for B-spline estimation is a trade-off between accuracy and complexity; higher values can also lead to over-fitting.

$\mapsto$ Select d and n to maximize the Bayesian information criterion (BIC):

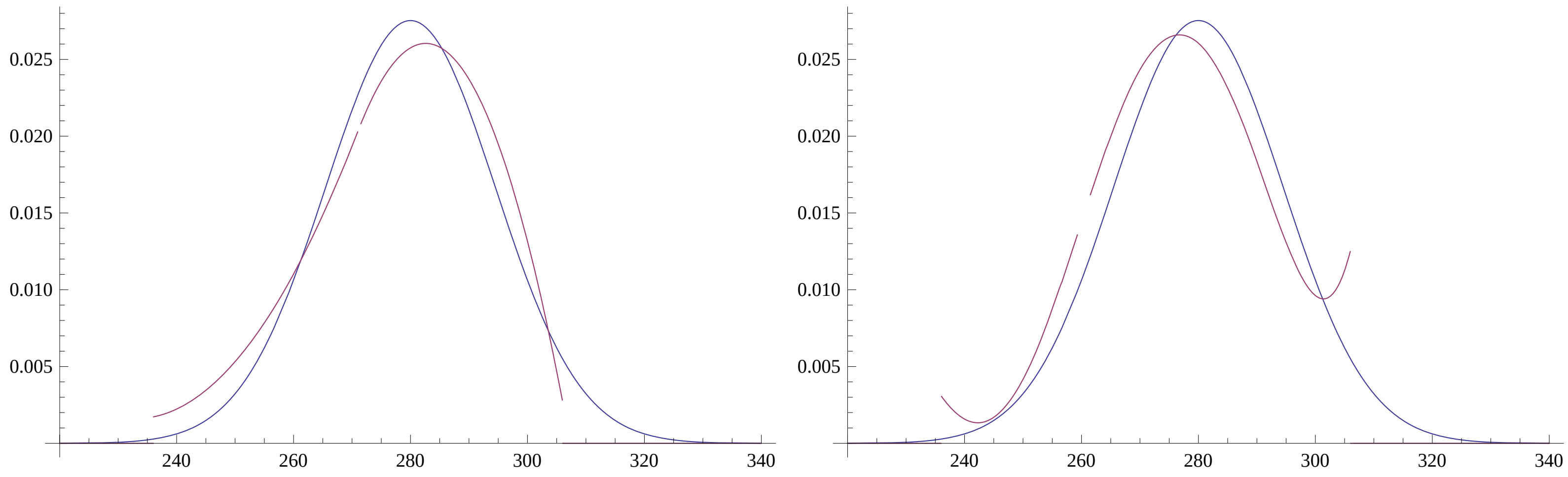
$$BIC\left(\hat{f}_{X|L=\ell}(x), \mathcal{D}\right) = \mathcal{L}\left(\mathcal{D}|\hat{f}_{X|L=\ell}(x)\right) - ((m-1)\log N)/2 \ .$$

The second term is a penalty for adding parameters to the model ($\mathcal{L}$ is the likelihood of the data given the model).

$\mapsto$ In practice, once we settle on d and n, we may not go through this step ($n = 2$ & $d = 3$ maximized BIC in this example).

Mixture of Polynomials (MOP)

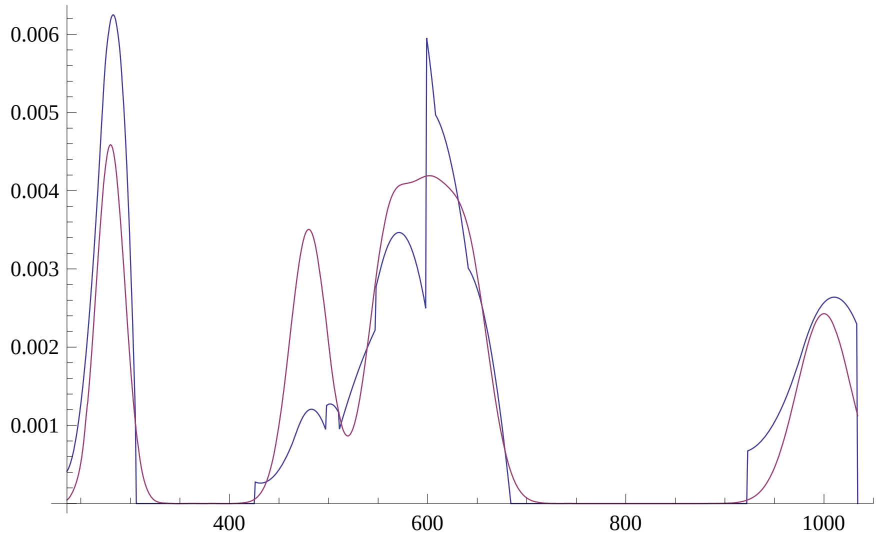
• What does "over-fitting" look like? Left: $n = 2$ & $d = 3$; Right: $n = 3$ & $d = 5$.



• Left: $\mathcal{L}\left(\mathcal{D}|\hat{f}_{L=7}\right) = -284.0$ & BIC: -290.4 ; Right: $\mathcal{L}\left(\mathcal{D}|\hat{f}_{L=7}\right) = -280.9$ & BIC: -293.7 (Bonus: model on left entails less computational complexity).

Mixture of Polynomials (MOP)

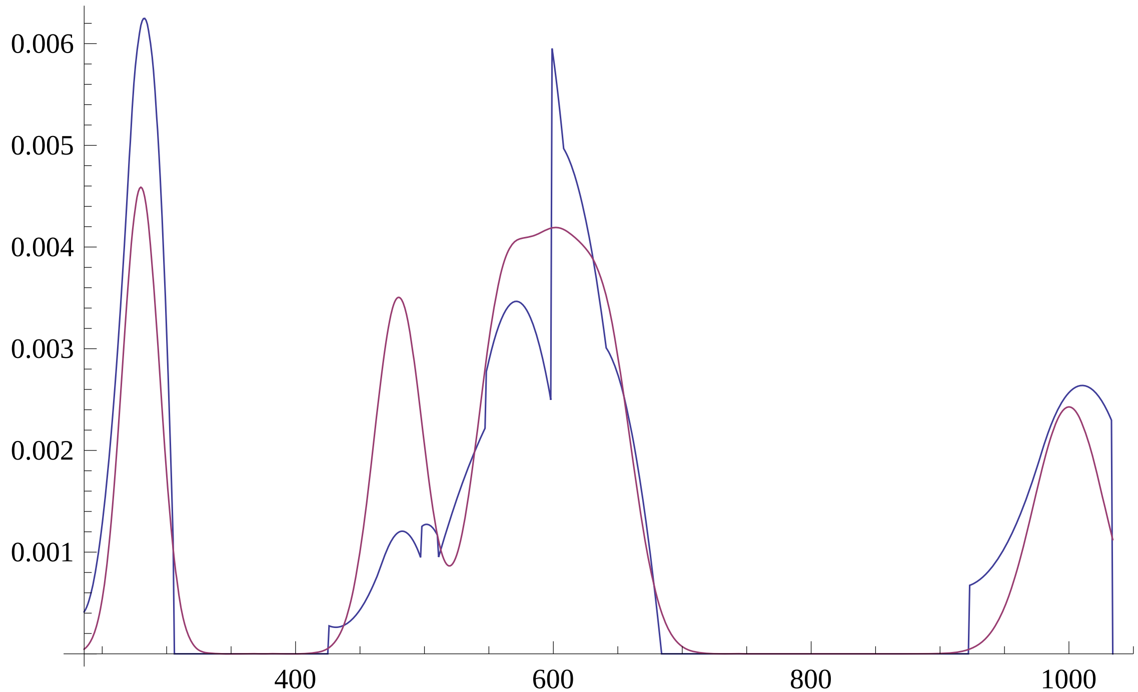
$\mapsto$ Fitting process repeated for $\ell = 12, 14, 15, 16, 25$.



The approximate LTD distribution:

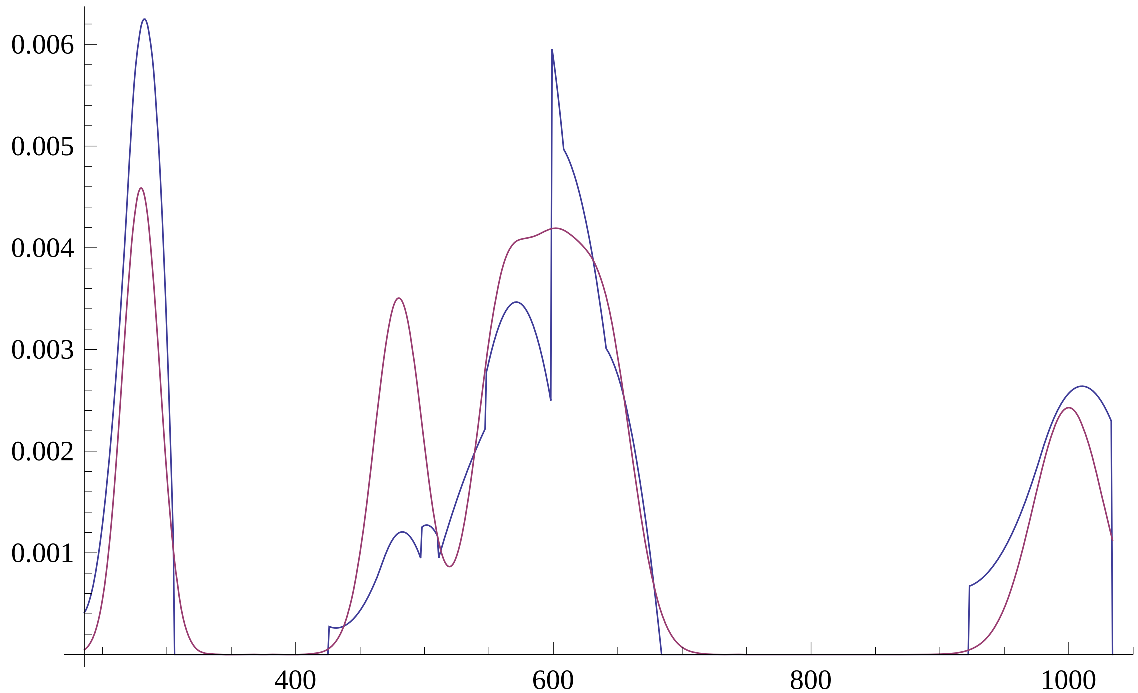
$$\widehat{f}_X(x) = \sum_{\ell^{(k)}=1}^{6} P\left(L = \ell^{(k)}\right) \cdot \widehat{f}_{X|L=\ell^{(k)}}(x)$$

Mixture of Polynomials (MOP)



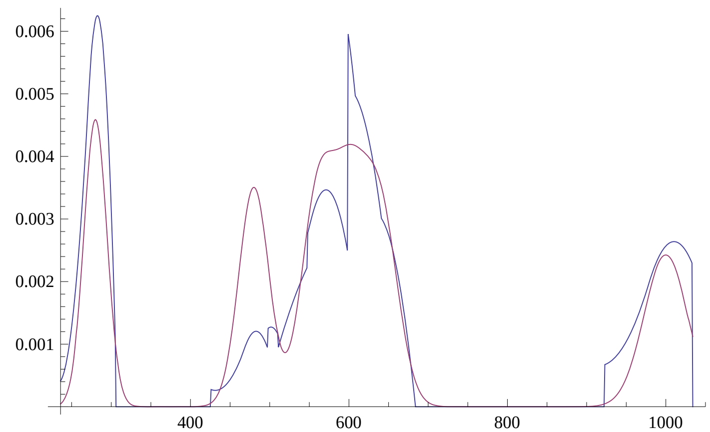
$\mapsto$ $\widehat{f}_X$ is overlaid on the mixture of normal distribution f_X (the "actual" distribution), but $\widehat{f}_X$ was not generated using knowledge of the underlying distribution.

Mixture of Polynomials (MOP)



$\mapsto \widehat{f}_X$ is relatively compact – it has 15 pieces and is a 2nd degree polynomial.

Mixture of Polynomials (MOP)



$\mapsto$ How is this useful?

Recall: Expected shortage per cycle: $\widehat{S}_R(R) = \int_R^{X_{max}} (x-R) \cdot \widehat{f}_X(x)\, dx$

$\mapsto$ $\widehat{S}_R$ can be calculated in closed-form (here a 10-piece, 5th degree MOP), so $\widehat{TC}_b$ is closed-form.

Finding Optimal Policies

Test case (from CH): $K_b = 50$, $K_s = 150$, $h_b = 5$, $h_s = 12.5$, $\pi = 6$, 250 working days so $Y = 250 * 40 = 10000$.

• Buyer would like to operate in a decentralized supply chain with $V = 0$ and set: $(Q_d^*, R_d^*) = \underset{(Q,R)}{\text{ArgMin}}\ \hat{TC}_b\,(Q, R, 0)$

$\mapsto$ Solution: $Q_d^* = 455$, $R_d^* = 1019$, $TC_b^d = \hat{TC}_b(Q_d^*, R_d^*, 0) = 4406.4$

• Given buyer's (Q_d^*, R_d^*), supplier finds $N_d^* = 1$ to minimize its costs.

$\mapsto TC_s^d = TC_s(Q_d^*, N_d^*, 0) = 3298$

$\mapsto TC^d = TC_b^d + TC_s^d = 4406.4 + 3298 = 7704.4$

Finding Optimal Policies

- Supplier would like to operate in a centralized supply chain

$\mapsto$ Define: $TC^c(Q,R,N) = TC_b(Q,R,0) + TC_s(Q,N,0)$

$$(Q_c^*, R_c^*, N_c^*) = \underset{(Q,R,N)}{\text{ArgMin}} \; TC^c(Q,R,N)$$

$\mapsto$ Solution: $Q_c^* = 909$, $R_c^* = 1004$, $N_c^* = 1$

$\mapsto$ $TC_b^c = 4955.9$, $TC_s^c = 1649.9$, $TC^c = 6605.8$

$\mapsto$ Coordination can save: $TC^+ = TC^d - TC^c = 7704.4 - 6605.8 = 1098.5$

Observations

- Buyer prefers a decentralized supply chain: $TC_b^d < TC_b^c$
- Supplier prefers a centralized supply chain: $TC_s^d > TC_s^c$

$\mapsto$ Centralized policy requires buyer to raise order quantity by $Q_c^* - Q_d^* = 454$ and can save $TC^+ = 1098.5$.

$\mapsto$ Centralized no. of orders: $Y/Q_c^* \approx 11$

$\mapsto$ Seller offers buyer rebate each cycle:

$$\overline{V} = 0.5 \cdot \frac{TC^+}{Y/Q_c^*} \approx 0.5 \cdot \frac{1098.5}{11} \approx 50$$

$\mapsto$ Parties can agree on other split of TC^+

Alternate Solution (Chaharsooghi and Heydari (2010) — CH)

- Approximate the density function f_X for X using a normal distribution using sample means and variances.

- For this example, $E(X) = \mu_X = 417.2$ and

$$Var(X) = \sigma_X^2 = 19881$$

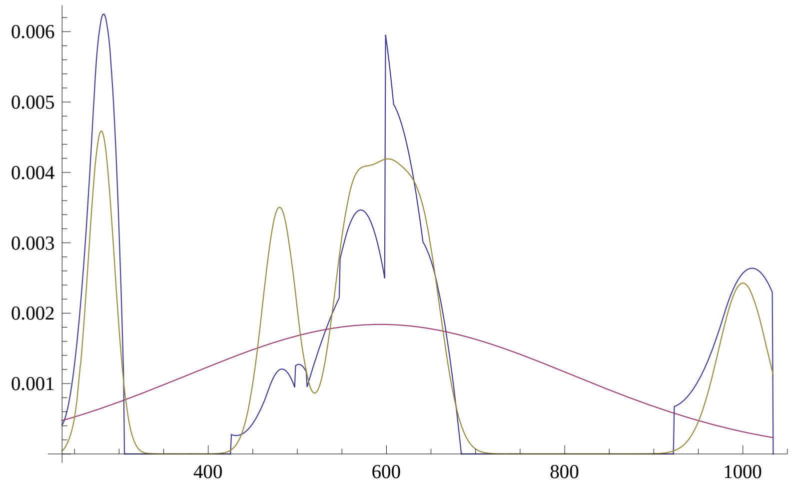
- Define: $k = (R - \mu_X)/\sigma_X$, so that

$$S(R) = S_k(k) \cdot \sigma_X = \sigma_X \cdot \int_k^{\infty} (z - k) \cdot \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \, dz \ .$$

- CH solve analytically for a partial solution for (Q, k, N) stated in terms of the cost parameters and normal CDF.

Alternatives – MOP and CH Solution

$\mapsto$ We should measure the effectiveness of the models by their "value in use"



• Consider two models: 1) MOP model; and 2) CH model. We will compare the solutions obtained from the two models by simulating from the (unknown) underlying "actual" normal daily demand and discrete LT distributions (the "actual" model).

Evaluating the Models

Decentralized	Q^*	R^*	N^*	Buyer TC	% Dec.	CPU (sec)
MOP (mixture dist.)	455	1019	1	4455	2.0%	1.77
CH (normal)	558	999	1	4531	0%	0.06

Coordinated	Q^*	R^*	N^*	SC TC	% Dec.	CPU (sec)
MOP (mixture dist.)	909	1004	1	6628	3.7%	2.16
CH (normal)	1012	926	1	6872	0%	0.09

• The costs are directly comparable — calculated by inserting the MOP and CH solutions into the simulation model and running 100,000 trials.

Air Force Example – F-15/16 Power Supply

• 1827 observations of daily demand (2008–2012) with $\overline{d} = 0.63$ and $s_d^2 = 1.09$.

$\mapsto$ Mode = 0 (1158 observations); Maximum = 8.

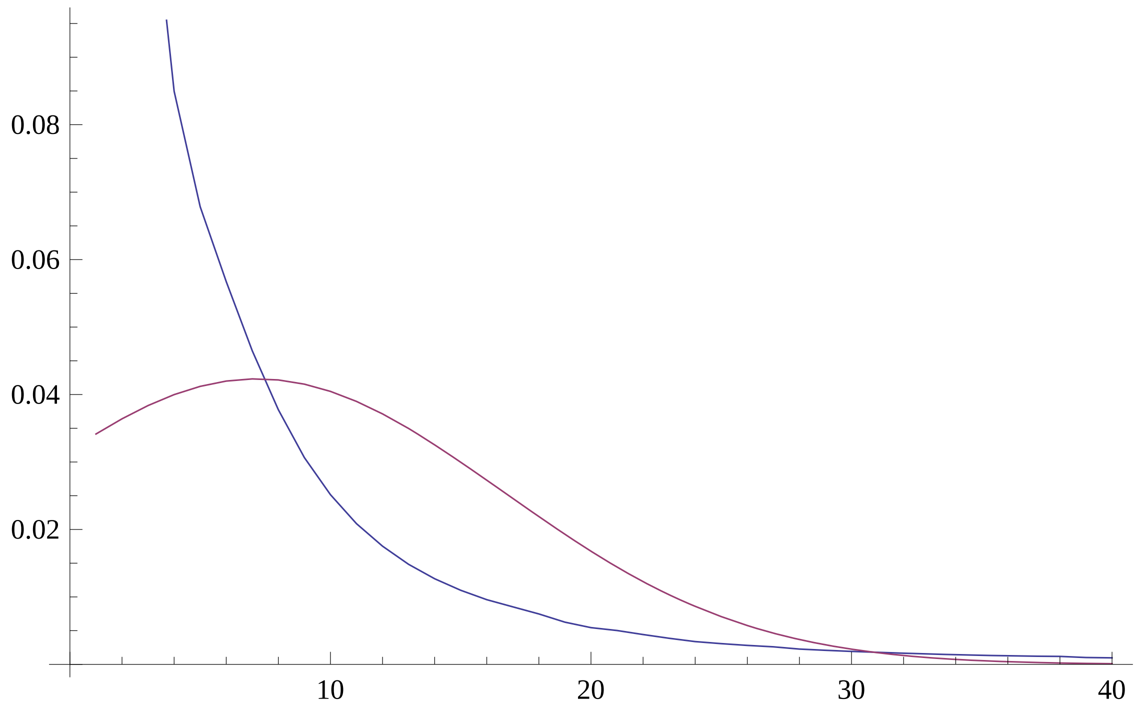
• 100 randomly sampled requisitions: $\overline{\ell} = 10.4$ and $s_\ell^2 = 162.8$ (Min=1; Max=73).

• Annual unit holding cost: 15%; Unit Price: \$224,392; h_b = \$33,658; K_b = \$5.20.

• Annual unit shortage cost (π) – if one unit short one officer at captain pay is 50% productive, π=25000.

Air Force Example

- LTD Distribution — MOP model and a normal approximation



- Use this distribution to find optimal Q^*, R^*, and N^* policies.

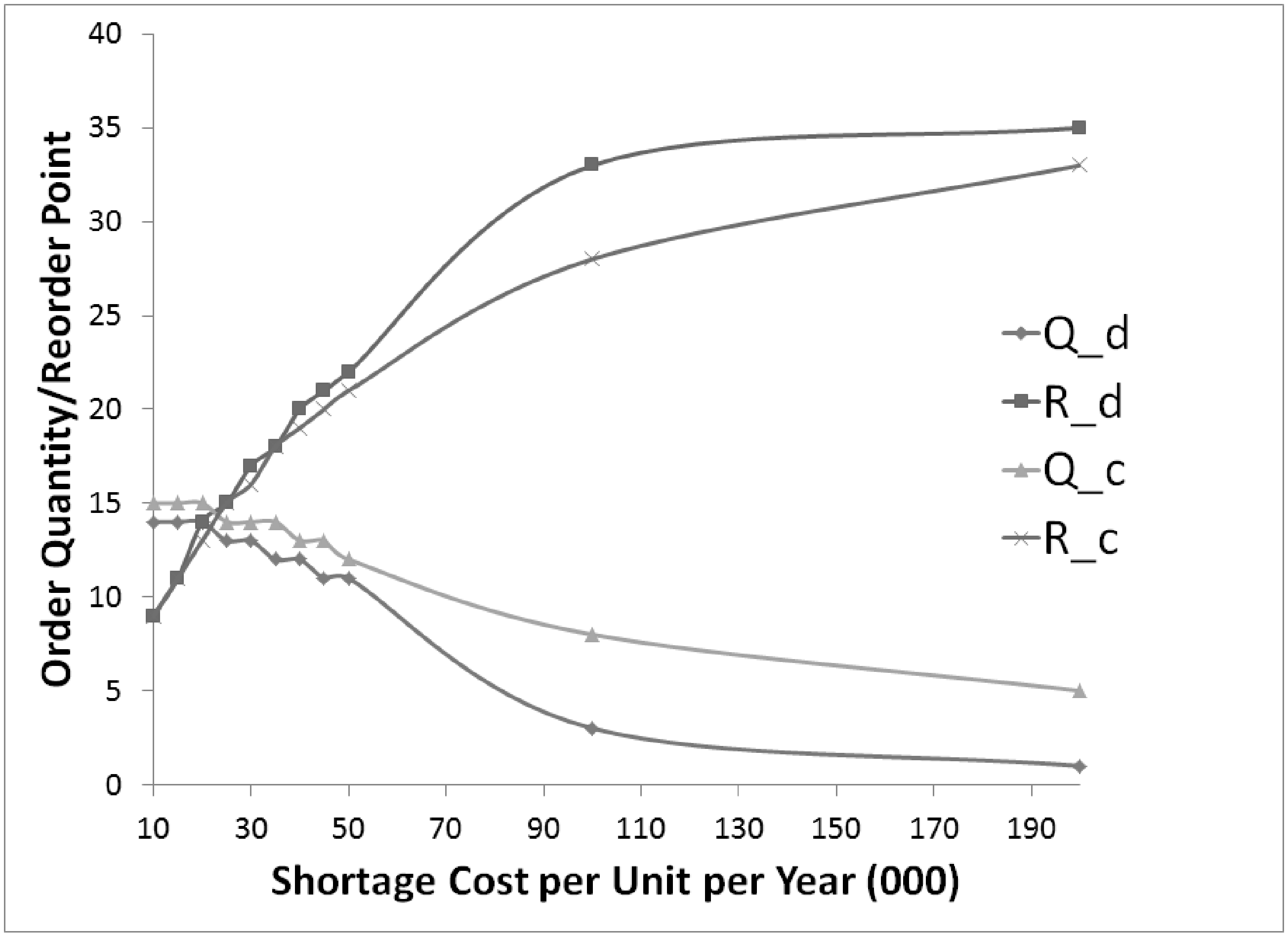
Air Force Example

- Comparison of solutions with MOP and normal approximations

					Holding	Order	Shortage	Total
		Q*	R*	N*	Cost	Cost	Cost	Cost
Decentralized	MOP	13	15	1	2,642,476	463	6,438	2,649,377
Coordinated	MOP	14	15	1	2,677,978	432	9,315	2,687,725
Decentralized	Normal	7	22	1	3,193,176	858	274	3,194,308
Coordinated	Normal	9	21	1	3,242,049	666	411	3,243,126

- Buyer costs calculated by implementing the policy with the actual demand data for 2008–2012 (before considering coordination incentives) and lead times drawn randomly for each order from the empirical distribution.

Sensitivity to Shortage Cost Parameter



Conclusions

• Mixture distributions can be used to model the distribution for demand during lead time using strictly empirical data with no limits on the underlying distribution.

• By using MOP distributions estimated from B-spline functions, we can perform integrations required to determine optimal order quantities, reorder points, and service levels in closed-form.

• Next steps: creating models under different sets of assumptions, e.g. a vendor-managed inventory model; improving the efficiency of the solution algorithm.